
MDIRA Bibliography Index Import File Specifications

Author: Peter Parker
Creation Date: Wednesday, July 14, 2004
Last Update: Tuesday, July 27, 2004
Version: 1.0 (Draft)



TABLE OF CONTENTS

TABLE OF CONTENTS	2
FORWARD.....	3
GENERAL FILE DETAILS.....	4
FILE FORMATS.....	5
COMMON ELEMENTS	5
NUMBERED SERIES ELEMENTS	7
ACADEMIC ELEMENTS	8
OTHER ELEMENTS	9
APPENDIX A	10

Forward

The reason for these specifications is to define the format of the import files used to batch load the system. There are three types of Publications that the system will currently accept references for. The types are: "Numbered Series", "Academic" and "Other". Since the number of data elements differs for each type, there are three different Import File formats for each type. Since all three references share similar elements, there is a section that describes the common elements and then three other sections that describe elements specific to the three Publication types.

General File Details

Files will be ASCII compatible. The file will be made up of data elements.

Each element will be separated by a pipe “|” symbol (shift + \ on the keyboard). Other delimiters will be described in the following sections when needed.

One reference or record must reside on one line of the file. A file may contain one or more references and each reference must be separated by a carriage return (EOL) character.

Each element of a reference may have special formatting requirements. These requirements will be detailed in the following sections. In the following sections, not null (must have a value) elements will be formatted in **red fonts** while elements that do not need a value will be formatted in **blue fonts**. All elements must be accounted for by a minimum of two pipe symbols with a space character between them (means no value). For an example file listing, see Appendix A.

As files are processed and loaded into the system, the Owner Agency element in the database is assumed to be the Agency of the user currently logged into the system.

File Formats

Common Elements

All import files will contain the common list of elements in the following order.

Publication_Type | **Title** | **Pubication_Year** |
Publication_Year_Type | **Release_Year** | **Edition** |
Larger_Work | **Pages** | **Sheets** | **Map_Scale** | **Repository** |
QMQs | **Authors** | **URLs** |

Where:

Publication_Type must be one of “NUMBERED_SERIES”, “ACADEMIC” or “OTHER”.

Title is freehand text that conforms to the STA.

Publication_Year is a four digit number that represents the publication year.

Publication_Year_Type is a text string and must be one of the following (“ACTUAL”¹, “DISCOVERED”, “ESTIMATED”).

Release_Year is a four digit number that represents the release year if different than the publication year above.

Edition is the edition number as an ordinal numeral or numerals. Formatting should be in compliance with that STA. Do not include the “ed.” or the beginning or end parentheses.

Larger_Work is a block of text with the contents of “*Editors* (same rule as Authors), *ed.s* (ed. or comp.s), *Title* (of the larger work), *Edition* (of the larger work)”. Include the “in” or “of” as necessary.

Pages is text that indicates the number of pages in publication. Text should include the post or pre fix of “p.” (See STA). This element may also contain the number of figures (figs) and or plates (pls). Formatting should conform to the STA.

Sheets is text that indicates the number of sheets if any. Text should include the post fix of “sheets” (see STA).

Map_Scale is the integer value of the denominator of the map scale if published as a map.

Repository is the physical location of the publication. This is not part of the STA and is only added to the end of the citation if available.

QMQs is the list of related USGS Quarter Million Quadrangles (QMQ) IDs. There must be at least one related QMQ. A list of valid Quadrangles can be found in the QMQ table. The

¹ ACTUAL is the default value.

DGGS_NUM column or the NAME can be used as the identifier. The list will be delimited by a semicolon (“;”).

Authors is the list of authors for the publication. There must be at least one author. The list will be delimited by a semicolon (“;”). The author’s name must be listed in compliance with that STA.

URLs is a list of online content (URLs) to web compliant pages or images of this publication. The list will be delimited by a semicolon (“;”).

Numbered Series Elements

Numbered Series type import files will contain the common elements plus the following:

Agency | **Series** | **Series_Number**

Where:

Agency is the Publishing Agency and is one of the AGENCY_ID or NAME values taken from the AGENCY table.

Series is a Series Abbreviation (SERIES_ABBRV) or NAME value taken from the PUBLISHING_AGENCY_SERIES table. The combination of the Agency ID (above) and the selected Series must be valid (exists in the table) before this element can be entered.

Series_Number is text and is the series number for this publication.

Complete File Format:

Publication_Type | Title | Publication_Year | Publication_Year_Type | Release_Year |
Edition | Larger_Work | Pages | Sheets | Map_Scale | Repository | QMQs | Authors | URLs |
Agency | Series | Series_Number

Total of 17 elements.

Academic Elements

Academic type import files will contain the common elements plus the following:

Publication_Place | **Institution** | **Degree**

Where:

Publication_Place is the text string indicating the place of publication.

Institution is the text string indicating the name of the institution that conferred the degree.

Degree is the academic degree level at which the report was written. This must be one of the following: "Ph.D. dissertation", "M.A. theses".

Complete File Format:

Publication_Type | **Title** | **Pubication_Year** | **Publication_Year_Type** | **Release_Year** |
Edition | **Larger_Work** | **Pages** | **Sheets** | **Map_Scale** | **Repository** | **QMQs** | **Authors** | **URLs** |
Publication_Place | **Institution** | **Degree**

Total of 17 elements.

Other Elements

For all other types of publications not described above include the common elements plus the following:

Other_Details

Data contained in Other Details is the text that is displayed in the citation for this reference that follows the colon (not including the colon) and precedes the Pages information. This data must conform to the STA, formatting is ultimately up to the Contributor.

Complete File Format:

Publication_Type | Title | Pubication_Year | Publication_Year_Type | Release_Year |
Edition | Larger_Work | Pages | Sheets | Map_Scale | Repository | QMQs | Authors | URLs |
Other_Details

Total of 15 elements.

Appendix A
